QUICKEST JOINT DETECTION AND CLASSIFICATION OF FAULTS IN STATISTICALLY PERIODIC PROCESSES

Taposh Banerjee⋆*, Smruti Padhy*†*, Ahmad Taha*⋆*, and Eugene John*⋆

⋆ Department of ECE, University of Texas at San Antonio
† Texas Advanced Computing Center, University of Texas at Austin

ABSTRACT

An algorithm is proposed to detect and classify a change in the distribution of a stochastic process that has periodic statistical behavior. The problem is posed in the framework of independent and periodically identically distributed (i.p.i.d.) processes, a recently introduced class of processes to model statistically periodic data. It is shown that the proposed algorithm is asymptotically optimal as the rate of false alarms and the probability of misclassification goes to zero. This problem has applications in anomaly detection in traffic data, social network data, ECG data, and neural data, where periodic statistical behavior has been observed. The effectiveness of the algorithm is demonstrated by application to real and simulated data.

***Index Terms*—** Quickest change detection and identification, ECG arrhythmia detection and identification, traffic congestion mode detection.

1. INTRODUCTION

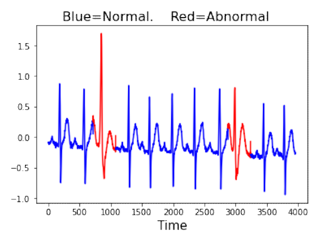
In the classical problem of quickest change detection [1], [2], [3], a decision maker observes a stochastic process with a given distribution. At some point in time, the distribution of the process changes. The problem objective is to detect this change in distribution as quickly as possible, with minimum possible delay, subject to a constraint on the rate of false alarms. This problem has applications in statistical process control [4], sensor networks [5], cyber-physical system monitoring [6], regime changes in neural data [7], traffic monitoring [8], and in general, anomaly detection [8], [9].

In many applications of anomaly detection, the observed process has periodic or regular statistical behavior (see Fig. 1). Some examples are as follows:

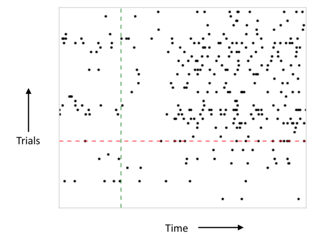
1. *Arrhythmia detection in ECG Data*: The electrocardiography (ECG) data has an almost periodic waveform pattern with a series of P waves, QRS complexes, and ST segments (see Fig. 1a). An arrhythmia can cause a change in this regular pattern [10].

2. *Detecting changes in neural spike data*: In certain brain-computer interface (BCI) studies [11], an identical experiment is performed on an animal in a series of trials leading to similar firing patterns. An event or a trigger (which is part of the experiment) can change the firing pattern after a certain trial (see Fig. 1b).

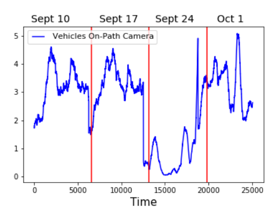
3. *Anomaly detection in city traffic data*: Vehicle counts at a street intersection in New York City (NYC) have been found to show regular patterns of busy and quiet periods [8,9,12–15]. A congestion or an accident can cause a drop or increase in these vehicle counts (see Fig. 1c).



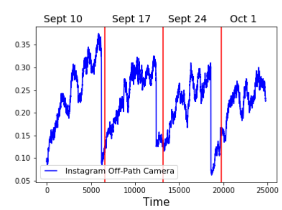
(a) ECG data from MIT-BIH database [10].



(b) Neural spike data from a mice BCI experiment [11].



(c) Vehicle counts observed through a CCTV camera in NYC [8].



(d) Instagram message counts posted near a CCTV camera in NYC [8]

Fig. 1: Real data from applications in medicine, neuroscience, and traffic exhibiting statistical periodicity.

4. *Social network data*: The count of the number of Instagram messages posted near a CCTV camera in NYC has also been found to show approximately periodic behavior [8,9,12–15] (see Fig. 1d).

5. *Congestion mode detection on highways*: In traffic density estimation problems, it is of interest to detect the mode (congested or uncongested) of the traffic first before deciding on a model to be used for estimation [16]. Motivated by the NYC data behavior, the traffic intensity in this application can also be modeled as statistically periodic.

In each of the above-mentioned applications, an event or anomaly can cause the statistical properties of the data to change. It is of interest to not only detect when this change occurs but also to classify the nature or type of change. In the ECG application, it is of interest to identify the nature or type of arrhythmia. In the BCI experiment, it is of interest to know the approximate time at which the firing pattern changes. In traffic application, we may be interested in knowing the location where the traffic is congested or an accident may have occurred. The anomaly detection problem in these applications can be posed as the problem of detecting deviations from this regular or pe-

The work of Taposh Banerjee was supported in part by the Cloud Technology Endowed Professorship IV. The work of Taposh Banerjee and Ahmad Taha was also partially supported by the National Science Foundation under Grant 1917164. The work of Taposh Banerjee and Eugene John was also partially supported by the National Science Foundation under Grant 2041327.

riodic statistical behavior. In this paper, we extend our work in [12] on quickest detection to joint quickest detection and classification.

At the outset, it may not be clear how to mathematically define statistical periodicity. This is discussed in Section 2.1. The precise joint change detection and classification problem is stated in Section 2.2 with the solution and optimality discussed in Section 3. In Section 4, we apply the algorithm to real and simulated data.

2. MODEL AND PROBLEM FORMULATION

2.1. Model for Statistical Periodicity

An independent and identically distributed (i.i.d.) process is a sequence of random variables that are mutually independent and have the same distribution. An independent and periodically identically distributed (i.p.i.d.) processes is defined as follows [12]:

Definition 1. *Let $\{X_n\}$ be a sequence of random variables such that the variable X_n has density f_n. The stochastic process $\{X_n\}$ is called independent and periodically identically distributed (i.p.i.d) if X_n are independent and there is a positive integer T such that the sequence of densities $\{f_n\}$ is periodic with period T:*

$$f_{n+T} = f_n, \quad \forall n \geq 1.$$

We say that the process is i.p.i.d. with the law $(f_1, \cdots, f_T)$.

The law of an i.p.i.d. process is completely characterized by the finite-dimensional product distribution involving $(f_1, \cdots, f_T)$. Our objective in this paper is to develop an algorithm that can observe the process $\{X_n\}$ in real-time and detect changes in the distribution as quickly as possible, subject to a constraint on the rate of false alarms. The proposed algorithm will also help classify or identify the post-change distribution with a low probability of misclassification.

2.2. Joint Detection and Classification Formulation

We assume that in a normal regime, the data can be modeled as an i.p.i.d. process with the law $(g_1^{(0)}, \cdots, g_T^{(0)})$. At some point in time ν, called the change point in the following, the law of the i.p.i.d. process is governed not by the densities $(g_1^{(0)}, \cdots, g_T^{(0)})$, but by one of the densities $(g_1^{(\ell)}, \cdots, g_T^{(\ell)})$, $\ell = 1, 2, \dots, M$, with

$$g_{n+T}^{(\ell)} = g_n^{(\ell)}, \quad \forall n \geq 1, \quad \ell = 1, 2, \dots, M.$$

Specifically, at the time point ν, the distribution of the random variables change from $\{g_n^{(0)}\}$ to $\{g_n^{(\ell)}\}$:

$$X_n \sim \begin{cases} g_n^{(0)}, & \forall n < \nu, \\ g_n^{(\ell)} & \forall n \geq \nu, \text{ for some } \ell = 1, 2, \dots, M. \end{cases} \tag{1}$$

We want to detect the change described in (1) as quickly as possible, subject to a constraints on the rate of false alarms and on the probability of misclassification. Mathematically, we are looking for a pair (τ, δ), where τ is stopping time, i.e.,

$$\{\tau \leq n\} \in \sigma(X_1, X_2, \dots, X_n),$$

and δ is a decision rule, i.e., a map such that

$$\delta(X_1, X_2, \dots, X_\tau) \in \{1, 2, \dots, M\}.$$

Let $\mathsf{P}_\nu^{(\ell)}$ denote the probability law of the process $\{X_n\}$ when the change occurs at time ν and the post-change law is $(g_1^{(\ell)}, \cdots, g_T^{(\ell)})$. We let $\mathsf{E}_\nu^{(\ell)}$ denote the corresponding expectation. When there is no change, we use the notation E_∞. The problem of interest is as follows [17]:

$$\begin{aligned} &\min_{\tau,\delta} \quad \max_{1\leq \ell \leq M} \sup_{\nu \geq 1} \operatorname{ess\,sup} \mathsf{E}_\nu^{(\ell)}[(\tau - \nu + 1)^+ | X_1, \cdots, X_{\nu-1}], \\ &\text{subj. to} \quad \mathsf{E}_\infty[\tau] \geq \beta, \\ &\text{and} \quad \mathsf{P}_1^{(\ell)}[\tau < \infty, \delta \neq \ell] \leq a_\beta \, \mathsf{E}_1^{(\ell)}[\tau], \quad \ell = 1, 2, \dots, M, \\ &\qquad \text{where } \log a_\beta^{-1} \sim \log \beta, \text{ as } \beta \to \infty. \end{aligned} \tag{2}$$

Here ess sup is used to denote the supremum of the random variable $\mathsf{E}_\nu^{(\ell)}[(\tau - \nu + 1)^+ | X_1, \cdots, X_{\nu-1}]$ outside a set of measure zero. Here and below, for two functions $h(\beta)$ and $f(\beta)$ of β, we use $f(\beta) \sim h(\beta)$, as $\beta \to \infty$, to denote that the ratio of the two functions goes to 1 in the limit. Further motivation of this and other problem formulations for change point detection and isolation can be found in the literature [2], [17], [18].

2.3. Algorithm for Detection when $M = 1$

When $M = 1$, i.e., when there is only one post-change i.p.i.d. law, then an algorithm that is asymptotically optimal for detecting a change in the distribution is the periodic-CUSUM algorithm proposed in [12]. In this algorithm, we compute the sequence of statistics

$$W_{n+1} = W_n^+ + \log \frac{g_{n+1}^{(1)}(X_{n+1})}{g_{n+1}^{(0)}(X_{n+1})} \tag{3}$$

and raise an alarm as soon as the statistic is above a threshold A:

$$\tau_c = \inf\{n \geq 1 : W_n \geq A\}. \tag{4}$$

Define

$$I_{10} = \frac{1}{T} \sum_{i=1}^{T} D(g_i^{(1)} \parallel g_i^{(0)}), \tag{5}$$

where $D(g_i^{(1)} \parallel g_i^{(0)})$ is the Kullback-Leibler divergence between the densities $g_i^{(1)}$ and $g_i^{(0)}$. Then, the following result is proved in [12].

Theorem 2.1 ([12]). *Let the information number I_{10} as defined in* (5) *satisfy $0 < I_{10} < \infty$. Then, with $A = \log \beta$,*

$$\mathsf{E}_\infty[\tau_c] \geq \beta,$$

and as $\beta \to \infty$,

$$\begin{aligned} &\sup_{\nu \geq 1} \ \textit{ess}\sup \mathsf{E}_\nu[(\tau_c - \nu + 1^+ | X_1, \cdots, X_{\nu-1}] \\ &\sim \inf_{\tau : \mathsf{E}_\infty[\tau] \geq \beta} \ \sup_{\nu \geq 1} \ \textit{ess}\sup \mathsf{E}_\nu[(\tau - \nu + 1)^+ | X_1, \cdots, X_{\nu-1}] \\ &\sim \frac{\log \beta}{I_{10}}. \end{aligned} \tag{6}$$

Thus, the periodic-CUSUM algorithm is asymptotically optimal for detecting a change in the distribtion, as the false alarm constraint $\beta \to 0$. Further, since the set of pre- and post-change densities $(g_1^{(0)}, \cdots, g_T^{(0)})$ and $(g_1^{(1)}, \cdots, g_T^{(1)})$ are finite, the recursion in (3) can be computed using finite memory needed to store these $2T$ densities.

3. ALGORITHM FOR JOINT DETECTION AND CLASSIFICATION

When the possible number of post-change distributions $M > 1$ and when we are also interested in accurately classifying the true post-change law, the periodic-CUSUM algorithm is not sufficient. We now propose an algorithm that can perform joint detection and classification.

For $\ell = 1, \dots, M$, define the stopping times

$$\tau_\ell = \inf\left\{n \geq 1 : \max_{1\leq k\leq n} \min_{0\leq m\leq M, m\neq \ell} \sum_{i=k}^{n} \log \frac{g_i^{(\ell)}(X_i)}{g_i^{(m)}(X_i)} \geq A\right\}. \tag{7}$$

The stopping time and decision rule for our detection-classification problem is defined as follows:

$$\begin{aligned} \tau_{dc} &= \min_{1\leq \ell\leq M} \tau_\ell, \\ \delta_{dc} &= \arg\min_{1\leq \ell\leq M} \tau_\ell. \end{aligned} \tag{8}$$

A window-limited version of the above algorithm is obtained by replacing each τ_ℓ in (7) by

$$\tilde{\tau}_\ell = \inf\left\{n : \max_{n-L_\beta\leq k\leq n} \min_{0\leq m\leq M, m\neq \ell} \sum_{i=k}^{n} \log \frac{g_i^{(\ell)}(X_i)}{g_i^{(m)}(X_i)} \geq A\right\}. \tag{9}$$

For $1 \leq \ell \leq M$ and $0 \leq m \leq M,\ m \neq \ell$, define

$$I_{\ell m} = \frac{1}{T}\sum_{i=1}^{T} D(g_i^{(\ell)} \parallel g_i^{(m)}), \tag{10}$$

and

$$I^* = \min_{1\leq \ell\leq M} \min_{0\leq m\leq M, m\neq \ell} I_{\ell m}. \tag{11}$$

Recall that we are looking for (τ, δ) such that

$$\mathsf{E}_\infty[\tau] \geq \beta(1 + o(1)), \text{ as } \beta \to \infty, \tag{12}$$

and

$$\begin{aligned} \mathsf{P}_1^{(\ell)}[\tau < \infty, \delta \neq \ell] \leq a_\beta\, \mathsf{E}_1^{(\ell)}[\tau], \quad \ell = 1, 2, \dots, M, \\ \text{where } \log a_\beta^{-1} \sim \log\beta, \text{ as } \beta \to \infty. \end{aligned} \tag{13}$$

Let

$$C_\beta = \{(\tau, \delta) : \text{conditions in (12) and (13) hold}\}. \tag{14}$$

Theorem 3.1. *Let the information number I^* be as defined in* (11) *and satisfy $0 < I^* < \infty$. Then, with $A = \log 4M\beta$,*

$$(\tau_{dc}, \delta_{dc}) \in C_\beta.$$

Also,

$$\begin{aligned} &\max_{1\leq \ell\leq M} \sup_{\nu\geq 1} \ \textit{ess}\sup \mathsf{E}_\nu^{(\ell)}[(\tau_{dc} - \nu + 1)^+ | X_1, \cdots, X_{\nu-1}] \\ &\sim \inf_{(\tau,\delta)\in C_\beta} \max_{1\leq \ell\leq M} \sup_{\nu\geq 1} \ \textit{ess}\sup \mathsf{E}_\nu^{(\ell)}[(\tau - \nu + 1)^+ | X_1, \cdots, X_{\nu-1}] \\ &\sim \frac{\log\beta}{I^*}, \ \text{as } \beta \to \infty. \end{aligned} \tag{15}$$

Finally, the window-limited version of the test (9) *also satisfies the same asymptotic optimality property as long as*

$$\lim \inf_{\beta\to\infty} \frac{L_\beta}{\log\beta} > \frac{1}{I^*}.$$

Proof. For $1 \leq \ell \leq M$ and $0 \leq m \leq M,\ m \neq \ell$, define

$$Z_i(\ell, m) = \log \frac{g_i^{(\ell)}(X_i)}{g_i^{(m)}(X_i)}$$

to be the log likelihood ratio at time i between the measures $\mathsf{P}_1^{(\ell)}$ and $\mathsf{P}_1^{(m)}$. In the rest of the proof, to write compact equations, we use $X_1^{\nu-1}$ to denote the vector

$$X_1^{\nu-1} = (X_1, X_2, \dots, X_{\nu-1}).$$

For each $1 \leq \ell \leq M$ and $0 \leq m \leq M,\ m \neq \ell$, we first show that the sequence $\{Z_i(\ell, m)\}$ satisfies the following statement:

$$\begin{aligned} &\sup_{\nu\geq 1} \operatorname{ess}\sup \mathsf{P}_\nu^{(\ell)} \left(\max_{t\leq n} \sum_{i=\nu}^{\nu+t} Z_i(\ell, m) \geq I_{\ell m}(1+\delta)n \,\middle|\, X_1^{\nu-1} \right) \\ &\qquad \to 0, \text{ as } n \to \infty, \quad \forall \delta > 0, \end{aligned} \tag{16}$$

where $I_{\ell m}$ is as defined in (10).

Towards proving (16), note that as $n \to \infty$

$$\frac{1}{n}\sum_{i=\nu}^{\nu+n} Z_i(\ell, m) \to I_{\ell m}, \quad \text{a.s. } \mathsf{P}_\nu^{(\ell)}, \ \forall \nu \geq 1. \tag{17}$$

The above display is true because of the i.p.i.d. nature of the observation process. This implies that as $n \to \infty$

$$\max_{t\leq n} \frac{1}{n}\sum_{i=\nu}^{\nu+t} Z_i(\ell, m) \to I_{\ell m}, \quad \text{a.s. } \mathsf{P}_\nu^{(\ell)}, \ \forall \nu \geq 1. \tag{18}$$

To show this, note that

$$\begin{aligned} &\max_{t\leq n} \frac{1}{n}\sum_{i=\nu}^{\nu+t} Z_i(\ell, m) \\ &\quad = \max\left\{ \max_{t\leq n-1} \frac{1}{n}\sum_{i=\nu}^{\nu+t} Z_i(\ell, m), \ \ \frac{1}{n}\sum_{i=\nu}^{\nu+n} Z_i(\ell, m) \right\}. \end{aligned} \tag{19}$$

For a fixed $\epsilon > 0$, because of (17), the LHS in (18) is greater than $I_{\ell m}(1-\epsilon)$ for n large enough. Also, let the maximum on the LHS be achieved at a point k_n, then

$$\max_{t\leq n} \frac{1}{n}\sum_{i=\nu}^{\nu+t} Z_i(\ell, m) = \frac{1}{n}\sum_{i=\nu}^{\nu+k_n} Z_i(\ell, m) = \frac{k_n}{n}\frac{1}{k_n}\sum_{i=\nu}^{\nu+k_n} Z_i(\ell, m).$$

Now k_n cannot be bounded because of the presence of n in the denominator. This implies $k_n > i$, for any fixed i, and $k_n \to \infty$. Thus, $\frac{1}{k_n}\sum_{i=\nu}^{\nu+k_n} Z_i(\ell, m) \to I_{\ell m}$. Since $k_n/n \leq 1$, we have that the LHS in (18) is less than $I_{\ell m}(1+\epsilon)$, for n large enough. This proves (18). To prove (16), note that due to the i.p.i.d. nature of the process

$$\begin{aligned} &\sup_{\nu\geq 1} \operatorname{ess}\sup \mathsf{P}_\nu^{(\ell)} \left(\max_{t\leq n} \sum_{i=\nu}^{\nu+t} Z_i(\ell, m) \geq I_{\ell m}(1+\delta)n \,\middle|\, X_1^{\nu-1} \right) \\ &\quad = \sup_{1\leq \nu\leq T} \mathsf{P}_\nu^{(\ell)} \left(\max_{t\leq n} \sum_{i=\nu}^{\nu+t} Z_i(\ell, m) \geq I_{\ell m}(1+\delta)n \right). \end{aligned} \tag{20}$$

The right hand side goes to zero because of (18) and because the maximum on the right hand side in (20) is over only finitely many terms.

Next, we show that the sequence $\{Z_i(\ell, m)\}$, for each $1 \leq \ell \leq M$ and $0 \leq m \leq M,\ m \neq \ell$, satisfies the following statement:

$$\lim_{n\to\infty} \sup_{k\geq\nu\geq 1} \operatorname{ess\,sup}\ \mathsf{P}_\nu^{(\ell)} \left(\frac{1}{n} \sum_{i=k}^{k+n} Z_i(\ell, m) \leq I_{\ell m} - \delta \;\middle|\; X_1^{\nu-1} \right) = 0, \quad \forall \delta > 0. \tag{21}$$

To prove (21), note that due to the i.p.i.d nature of the process we have

$$\begin{aligned} &\sup_{k\geq\nu\geq 1} \operatorname{ess\,sup}\ \mathsf{P}_\nu^{(\ell)} \left(\frac{1}{n} \sum_{i=k}^{k+n} Z_i(\ell, m) \leq I_{\ell m} - \delta \;\middle|\; X_1^{\nu-1} \right) \\ &= \sup_{\nu+T\geq k\geq\nu\geq 1} \mathsf{P}_\nu^{(\ell)} \left(\frac{1}{n} \sum_{i=k}^{k+n} Z_i(\ell, m) \leq I_{\ell m} - \delta \right) \\ &= \max_{1\leq\nu\leq T} \max_{\nu\leq k\leq\nu+T} \mathsf{P}_\nu^{(\ell)} \left(\frac{1}{n} \sum_{i=k}^{k+n} Z_i(\ell, m) \leq I_{\ell m} - \delta \right). \end{aligned} \tag{22}$$

The right hand side of the above equation goes to zero for any δ because of (17) and also because of the finite number of maximizations. The theorem now follows from Theorem 4 of [17]. ■

Remark 1. *The algorithm and optimality are valid for multistream data as well. We only need to treat the observation process as a sequence of random vectors.*

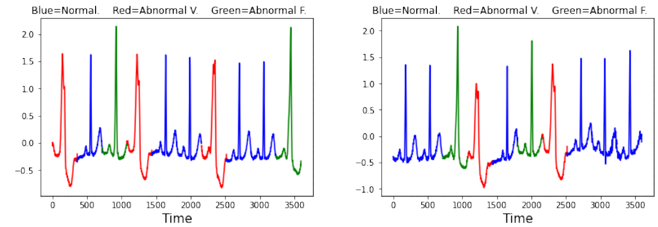
Remark 2. *For faster computation, in the numerical results section below, we use a recursive version of the algorithm by swapping the max and min operations in* (7)*.*

4. APPLICATIONS TO REAL AND SIMULATED DATA

We first apply our algorithm to ECG data from MIT-BIH database [10] record number 208 (see Fig. 2). The patient has two types of arrhythmias: premature ventricular contraction (V) and the fusion of ventricular and normal beat (F). Gaussian i.p.i.d. processes (joint Gaussian in each period) were trained using data from each class. In Fig. 2a and Fig. 2b, samples of ten waveforms from two different parts of the record are plotted. The normal ECG beats are shown in blue, the beats with arrhythmia of type V are shown in red, and the ones with arrhythmia of type F are shown in green. For each waveform, we computed the statistic

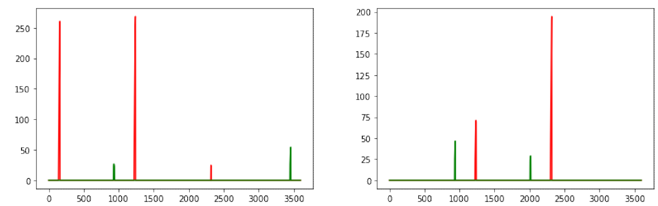
$$\min_{0\leq m\leq M, m\neq\ell} \quad \max_{1\leq k\leq n} \sum_{i=k}^{n} \log \frac{g_i^{(\ell)}(X_i)}{g_i^{(m)}(X_i)} \tag{23}$$

for each class (arrhythmia). The corresponding statistics are in Fig. 2c and Fig. 2d, respectively. For better visualization, the statistics were only computed for time slots in which the training data showed any disparity between data from different classes resulting in spiking behavior of the statistic. Moreover, all statistic values below a certain level were reset to zero. See also Remark 2. The statistics for class V are plotted in red and those for class F are plotted in green. The figures show that when there is an arrhythmia of type V (respectively, F) in the ECG data, there is a spike in the red (respectively, green) statistics. When the ECG beat is normal, there



(a) ECG data for patient 208 starting beat number 1100.

(b) ECG data for patient 208 starting beat number 1500.



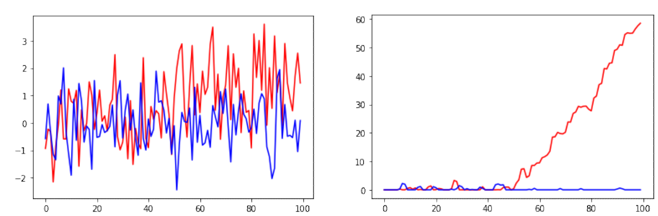
(c) Statistic (23) for data in Fig. 2a above.

(d) Statistic (23) for data in Fig. 2b above.

Fig. 2: Application of the proposed algorithm to MIT-BIH ECG data [10]. See also Remark 2. The spiking nature of the statistic is explained below (23).

is no spike in either of the statistics. A more comprehensive discussion on this application will be reported in a detailed version of this paper.

Next, we apply the proposed algorithm to multistream data (see Fig. 3 and Remark 1 above). We use two streams. For each stream, the data is i.i.d. $\mathcal{N}(0, 1)$ before change and alternates between $\mathcal{N}(2, 1)$ and $\mathcal{N}(0.5, 1)$ after change. At the change point (time 50 in Fig. 3a), the distribution in only one stream is affected. We again compute the statistic (23) for each stream. The data and the corresponding statistics are shown in the same color. In the example taken here, the red stream is affected post-change. This is reflected in the change in the drift of the red statistic in Fig. 3b after the change point.



(a) Multistream i.p.i.d. data for two channels. The period of the i.p.i.d. data is $T = 2$.

(b) Statistic (23) computed for data from each channel in Fig. 3a.

Fig. 3: Application of the proposed algorithm to simulated multistream data. See also Remark 1.

5. CONCLUSIONS AND FUTURE WORK

We have developed an asymptotic theory for joint quickest detection and classification (identification) of changes in i.p.i.d. models. In our future work, we will thoroughly investigate the application of the proposed algorithm and its variations to ECG data. We will also apply the algorithm to traffic data to detect and isolate anomalies or congestion modes.

6. REFERENCES

[1] H. V. Poor and O. Hadjiliadis, *Quickest detection*. Cambridge University Press, 2009.

[2] A. G. Tartakovsky, I. V. Nikiforov, and M. Basseville, *Sequential Analysis: Hypothesis Testing and Change-Point Detection*. Statistics, CRC Press, 2014.

[3] V. V. Veeravalli and T. Banerjee, *Quickest Change Detection*. Academic Press Library in Signal Processing: Volume 3 – Array and Statistical Signal Processing, 2014. `http://arxiv.org/abs/1210.5552`.

[4] G. Tagaras, "A survey of recent developments in the design of adaptive control charts," *Journal of Quality Technology*, vol. 30, pp. 212–231, July 1998.

[5] T. Banerjee and V. V. Veeravalli, "Data-efficient quickest change detection in sensor networks," *IEEE Transactions on Signal Processing*, vol. 63, no. 14, pp. 3727–3735, 2015.

[6] Y. C. Chen, T. Banerjee, A. D. Domnguez-Garca, and V. V. Veeravalli, "Quickest line outage detection and identification," *IEEE Transactions on Power Systems*, vol. 31, pp. 749–758, Jan 2016.

[7] T. Banerjee, S. Allsop, K. M. Tye, D. Ba, and V. Tarokh, "Sequential detection of regime changes in neural data," *arXiv preprint arXiv:1809.00358*, 2018.

[8] T. Banerjee, G. Whipps, P. Gurram, and V. Tarokh, "Sequential event detection using multimodal data in nonstationary environments," in *Proc. of the 21st International Conference on Information Fusion*, July 2018.

[9] T. Banerjee, G. Whipps, P. Gurram, and V. Tarokh, "Cyclostationary statistical models and algorithms for anomaly detection using multi-modal data," in *Proc. of the 6th IEEE Global Conference on Signal and Information Processing*, Nov. 2018.

[10] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals," *circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[11] Y. Zhang, N. M. Shinitski, S. Allsop, K. Tye, and D. Ba, "A two-dimensional seperable random field model of within and cross-trial neural spiking dynamics," *Neural Computation*, vol. 15, no. 5, pp. 965–991, 2003.

[12] T. Banerjee, P. Gurram, and G. Whipps, "Quickest detection of deviations from periodic statistical behavior," in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 5351–5355, IEEE, 2019.

[13] T. Banerjee, P. Gurram, and G. Whipps, "Bayesian quickest detection of changes in statistically periodic processes," in *2019 IEEE International Symposium on Information Theory (ISIT)*, pp. 2204–2208, IEEE, 2019.

[14] T. Banerjee, P. Gurram, and G. Whipps, "A sequential detection theory for statistically periodic random processes," in *2019 57th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 290–297, IEEE, 2019.

[15] T. Banerjee, P. Gurram, and G. Whipps, "Multislot and multistream quickest change detection in statistically periodic processes," in *2020 IEEE International Symposium on Information Theory (ISIT)*, pp. 1147–1152, IEEE, 2020.

[16] S. C. Vishnoi, S. A. Nugroho, A. F. Taha, C. Claudel, and T. Banerjee, "Asymmetric cell transmission model-based, ramp-connected robust traffic density estimation under bounded disturbances," in *2020 American Control Conference (ACC)*, pp. 1197–1202, IEEE, 2020.

[17] T. L. Lai, "Sequential multiple hypothesis testing and efficient fault detection-isolation in stochastic systems," *IEEE Transactions on Information Theory*, vol. 46, no. 2, pp. 595–608, 2000.

[18] I. V. Nikiforov, "A lower bound for the detection/isolation delay in a class of sequential tests," *IEEE Trans. Inf. Theory*, vol. 49, pp. 3037–3046, Nov. 2003.